

2012

Serial recall, word frequency, and mixed lists: the influence of item arrangement

Leonie M. Miller

University of Wollongong, leoniem@uow.edu.au

Steven J. Roodenrys

University of Wollongong, steven@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/sspapers>



Part of the [Education Commons](#), and the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Miller, Leonie M. and Roodenrys, Steven J., "Serial recall, word frequency, and mixed lists: the influence of item arrangement" (2012). *Faculty of Social Sciences - Papers*. 140.
<https://ro.uow.edu.au/sspapers/140>

Serial recall, word frequency, and mixed lists: the influence of item arrangement

Abstract

Studies of the effect of word frequency in the serial recall task show that lists of high-frequency words are better recalled than lists of low-frequency words; however, when high- and low-frequency words are alternated within a list, there is no difference in the level of recall for the two types of words, and recall is intermediate between lists of pure frequency. This pattern has been argued to arise from the development of a network of activated long-term representations of list items that support the redintegration of all list items in a nondirectional and nonspecific way. More recently, it has been proposed that the frequency effect might be a product of the coarticulation of items at word boundaries and their influence on rehearsal rather than a consequence of memory representations. The current work examines recall performance in mixed lists of an equal number of high- and low-frequency items arranged in contiguous segments (i.e., HHHLLL and LLLHHH), under quiet and articulatory suppression conditions, to test whether the effect is (a) nondirectional and (b) dependent on articulatory processes. These experiments demonstrate that neither explanation is satisfactory, although the results suggest that the effect is mnemonic. A language-based approach to short-term memory is favored with emphasis on the role of speech production processes at output.

Keywords

mixed, lists, influence, recall, serial, frequency, item, word, arrangement

Disciplines

Education | Social and Behavioral Sciences

Publication Details

Miller, L. M. & Roodenrys, S. J. (2012). Serial recall, word frequency, and mixed lists: the influence of item arrangement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38 (6), 1731-1740.

Serial Recall, Word Frequency and Mixed Lists: The Influence of Item Arrangement

Leonie M. Miller and Steven Roodenrys

University of Wollongong

Wollongong, Australia

Author Note

Leonie M. Miller, School of Psychology, University of Wollongong; Steven Roodenrys, School of Psychology, University of Wollongong.

A portion of this research (Experiment 1) was conducted as part of the first author's doctoral thesis and was funded by an Australian Postgraduate Award.

Corresponding author: Leonie M. Miller, School of Psychology, University of Wollongong, Wollongong NSW 2522 Australia; Telephone number: +61 2 4221 4454; Email: leoniem@uow.edu.au

Abstract

Studies of the effect of word frequency in the serial recall task show that lists comprised of high-frequency words are better recalled than lists of low-frequency words, however, when high and low frequency words are alternated within a list there is no difference in the level of recall for the two types of words, and recall is intermediate between lists of pure frequency. This pattern has been argued to arise from the development of a network of activated long-term representations of list items that support the reintegration of all list items in a non-directional and non-specific way. More recently, it has been proposed that the frequency effect might be a product of the coarticulation of items at word boundaries, and their influence on rehearsal, rather than a consequence of memory representations. The current work examines recall performance in mixed lists comprising an equal number of high and low-frequency items arranged in contiguous segments (i.e. HHHLLL and LLLHHH), under quiet and articulatory suppression conditions, to test whether the effect is (i) non-directional and (ii) dependent on articulatory processes. These experiments demonstrate that neither explanation is satisfactory although the results suggest that the effect is mnemonic. A language-based approach to short-term memory is favored with emphasis on the role of speech production processes at output.

Short-term serial recall tasks involve the auditory or visual presentation of a short list of items (typically 5-7 words or nonwords) immediately followed by an attempt to recall the items in their presentation order. Such tasks are presumed to involve the order-sensitive maintenance of a series of phonological representations upon which correct recall is dependent (Hulme et al., 1997; Baddeley, 1986). Manipulation of the attributes of verbal items within this paradigm assesses the impact those attributes have on the processes argued to underpin the task. One attribute, word frequency, is the most studied variable of language items in short-term memory (STM, Morin, Poirier, Fortin, & Hulme, 2006). Measures of word frequency estimate how often a word is encountered in written or spoken language and thus are an index of experience with verbal items.

The frequency effect in short-term serial recall has proven elusive to define, primarily because the manifestation of the effect is confounded with how frequency is manipulated within trials. Much of the early work on the frequency effect focused on the use of pure lists, when all list items are either high or low frequency (LF). Under this manipulation a robust and replicable advantage for high frequency (HF) words is observed (e.g., Allen & Hulme, 2006; Gregg, Freedman, & Smith, 1989; Hulme et al., 1997; Hulme, Stuart, Brown, & Morin, 2003; Poirier & Saint-Aubin, 1996; Tehan & Humphreys, 1988; Watkins & Watkins, 1977; Woodward, Macken, & Jones, 2008). The effect was presumed to reflect the operation of an item-specific property on recall processes. For example, the redintegration hypothesis (Hulme et al., 1997) argued that word frequency indexed the accessibility of long-term knowledge for each item. Adopting a two store approach to memory, it was argued that short-term phonological representations of items would become degraded within the life of a trial, and require the assistance of long-term phonological representations to restore them at the point of retrieval (Hulme et al., 1997; Poirier & Saint-Aubin, 1996; Roodenrys, Hulme, Alban, Ellis, & Brown, 1994; Schweickert, 1993). Specifically, a partial short-term trace, retrieved

from the phonological short-term store, would be restored by matching it against suitable long-term memory (LTM) candidates, in much the same way that a speech error might be corrected. The better recall of HF than LF items would result from the greater efficiency and reliability of access to long-term phonological representations for these words. Furthermore, as word frequency did not alter the rate of order errors in the recall of pure lists (Hulme et al., 2003; Poirier & Saint-Aubin, 1996), an explanation that located its influence after the selection of a trace for output, namely at a late-stage in the retrieval process, seemed likely.

However, when HF and LF items are alternated within the same list a different picture of the frequency effect emerges. While order memory remains unaffected by the frequency of items (Hulme et al., 2003), HF and LF words are recalled at similar levels, and the global performance on these lists is intermediate with respect to recall performances of pure HF and pure LF lists. An item-specific redintegration approach predicts that HF and LF items are recalled equally well in pure and alternating list formats – serial recall for alternating lists should resemble a sawtooth pattern enveloped by the serial position curves for pure lists. However the empirical evidence clearly refutes this (Hulme et al., 2003; Morin et al., 2006).

These results suggest that the nature of the frequency effect in serial recall is context dependent; how well a HF or LF word is recalled is a function of the list in which it is presented, not just a function of the word *per se*. To extend the redintegration hypothesis, a list of presented items might form a network of activation in LTM that supports late-stage reconstruction in a nonspecific way (Hulme et al., 2003; Stuart & Hulme, 2000; 2009). When a word is presented as part of a list, this activation will spread to other members of the list cohort through the pathways connecting them in LTM. Furthermore, the level of spreading activation to other list members will be governed by the strength of association between items as a function of everyday language use (Deese, 1959, 1960). High frequency items in pure lists would be assisted by high levels of activation in LTM, as HF items share strong pre-

existing associations due to their high rate of co-occurrence in natural language. In contrast LF in pure lists would be supported the least well, as LF items share weak associations resulting from low co-occurrence in everyday language use. In the context of alternating lists, HF and LF items would be supported by an intermediate level of long-term activation as association strengths between HF and LF items would be intermediate. This is due to the tendency for HF words to be well associated with other words in language, and so across all possible associations between the HF and LF items in a list, the general level of activation would be intermediate (Hulme et al., 2003; Stuart & Hulme, 2000). Importantly, this explanation implies that the level of supporting activation, determined by the list cohort, is non-directional and equally available to all list items.

A second complication in interpreting the frequency effect in the serial recall task involves frequency-based differences in articulation and the role of articulation in the task more generally. High frequency words are articulated faster than LF words of the same length and early explanations of the effect argued that rehearsal rate differences were responsible for the better retention of HF words (Wright, 1979). However, other experiments that accounted for speech rate differences, either in terms of experimental design (Gregg et al., 1989; Tehan & Humphreys, 1988) or statistical adjustment (Roodenrys et al., 1994; Hulme et al., 1997; Hulme et al., 2003), did not eliminate a frequency effect, suggesting that the source of the effect was not due to articulatory factors.

More recently, Woodward et al. (2008) proposed that the frequency effect in serial recall could be explained in terms of differences in the familiarity of articulation of items at word boundaries. They argued that HF words might appear in more articulatory contexts than LF words resulting in the better negotiation of the articulatory transitions between words in the pure lists used in STM tasks. They gave participants a serial recall task using the short HF and LF stimuli of Hulme et al. (1997) that was followed by the measurement of articulatory

duration of single items, pairs of items and sequences of 6 items using a reading aloud task. Measures of single and pairwise item duration involved the length of time taken to perform 10 repetitions. The serial recall task produced a frequency effect with HF items better recalled than LF items. The articulation duration data showed that HF words were produced reliably better than LF items, and all forms of duration measure revealed frequency effects. However, the effect for single item duration was less than for the pair or six-item sequences. As earlier work had not employed list-length sequences in measuring articulation rate, the influence of this factor may well have been underestimated.

Woodward et al. (2008) argued that insufficient consideration of the salience of the sequence-level differences in articulation duration had been given in previous studies of word frequency. If fluency across word boundaries was responsible for the greater proportion of articulation duration differences between pure HF and LF lists, then it was likely to impact the efficiency of rehearsal in serial recall. According to this view, rehearsal supports the development of a speech output plan to be used during the output phase; rehearsal emulates the movements of the vocal tract and this process converts a sequence of individual items into a single utterance (Hughes et al., 2009; Woodward et al., 2008). Conditions that present more difficult articulatory transitions between items, for example lists constructed from LF words, pose a greater challenge to the assembly of articulatory gestures that form the output stream, on which recall is based.

An extension of the co-articulation argument to lists of alternating frequency would suggest that familiarity with word boundaries between HF and LF words is intermediate given that, due to the greater numbers of articulatory contexts in which HF words occur, some experience in the transitions between HF and LF words would be typical. If so, this level of familiarity should translate to a speech output plan that is more efficiently produced than for pure LF lists, but less efficiently produced than for HF lists, resulting in recall

performance that is intermediate for all items in the list. Furthermore, assuming equivalence in the levels of familiarity of boundaries straddling the transition from HF to LF words, and LF to HF words respectively, it would be anticipated that HF and LF words would be recalled equally well. Therefore, this account is capable of explaining the empirical results for alternating lists observed by Hulme et al. (2003).

In the current research we sought to test the proposition made by Hulme et al. (2003) that the nature of the frequency effect in mixed lists of HF and LF words is non-directional. Hulme et al. (2003) had noted the possibility that associativity between adjacent list items, rather than inter-item associativity across all list items, could also explain the empirical observations made with alternating lists. They referred to the Temporal Context Model (TCM, Howard & Kahana, 2002a; 2002b), drawn from the free recall literature, as a model specifying a mechanism that links the semantic association between adjacent list items and recall performance.

In this model, recall is dependent on reinstating the temporal context of an item, and the recall of each item is used as a cue to retrieve the next temporal context. At presentation, a representation of context for an item is formed using a slowly evolving context state and other input based on the previous contexts in which the item has been experienced. The slow changing nature of the contextual signal means that items presented close together will have similar contextual representations. Furthermore, if the previous episodes in which adjacent items were experienced are closer together in time, as would be likely for semantically associated terms in language, then the previous temporal contexts that contribute to the formation of the current contexts also would be similar. Therefore, adjacent list items that possess common contextual history will produce cues better able to successfully drive the recall process.

Hulme et al. (2003) conceded that the compatibility of the TCM with the frequency effect in serial recall was not assured, as it was unknown whether the model could accommodate an effect that operated on item information, but not order information, and more recently, the TCM has also been found to contain a number of weaknesses from a computational perspective (see Farrell & Lewandowsky, 2008). Nonetheless, at a conceptual level, the possibility of pre-existing associations between adjacent items affecting serial recall performance remains a valid one (Hulme et al., 2003; Stuart & Hulme, 2000, 2009).

If the frequency effect is non-directional as Hulme et al. (2003) contend, then all lists containing equal numbers of HF and LF items should produce recall similar to those found for alternating lists, even if they differ in item arrangement. Experiment 1 directly compared the recall of pure lists with lists where each half is composed of items of one frequency type (i.e. HHHLLL or LLLHHH) and sought to test whether a directional component of the frequency effect is evident.

The co-articulation hypothesis (Woodward et al., 2008) would predict that a directionally sensitive frequency effect arises from this mixed list format, under the assumption that differences in the co-articulation of contiguously presented HF and LF items occur. Specifically, if the co-articulation of a series of 3 HF items at the beginning of a half list is more efficient than the co-articulation of 3 LF items, then it can be anticipated that a frequency effect, particularly in the first half of the list will eventuate, as the early list items are more likely to be affected by rehearsal (Page & Norris, 1998). This pattern would stand in contrast to the recall of alternating lists, where presumably co-articulation effects would be similar regardless of whether the list commenced with a HF or an LF word.

Method

Participants

Thirty-six undergraduate and postgraduate University of Wollongong students (24 females, 12 males) participated in this experiment for course credit. Participants were native Australian-English speakers. The mean age of the participants whose data was 23.1 years ($SD = 7.6$ years).

Materials

Two stimulus sets of 96 CVC words were created based on the manipulation of word frequency. Frequency ratings were taken from the Celex database (Baayen, Piepenbrock, & Van Rijn, 1993) and were the composite of database entries for the same orthography and across those entries with the same word identification number, so that for example, the frequency counts for *bird* and *birds* were combined. Additionally, the measures of homophones were included in the calculations of item attributes, where applicable.

High frequency words were shown to be greater in \log_{10} per million frequency ratings ($M = 2.21$, $SD = 0.27$) than LF words ($M = 0.70$, $SD = 0.36$), *Mann-Whitney* $U = 0.00$, $p < .001$. These sets were matched phonological neighborhood size (derived from the Celex database), concreteness ratings (Coltheart, 1981), and a series phonological similarity measures (PSIMETRICA, Mueller, Seymour, Kieras & Meyer, 2003) for onset, nucleus, and coda components of words. The sets were also matched on the number of phonological neighbors of each item within their respective word sets. The number of words with short, long or diphthong vowels for LF and HF sets identified that there was a trend towards the LF set containing more words with short vowels.¹

Given the possibility that the co-articulation at word boundaries might impact on the results of the present experiments, the stimulus sets were subjected to further examination

¹ The stimuli and a detailed description of the selection process are available from the corresponding author on request.

regarding the characteristics of the word boundaries in each word set. An analysis of the articulatory complexity for all coda-onset combinations in each set was conducted (A. Murray & Jones, 2002). As each set contained 96 items, 9120 word boundaries were assessed in terms of voicing, manner of articulation and place of articulation. The results of this examination are presented in Table 1, where each figure represents the number of transitions of each type.

Table 1 demonstrates that in this case, phonetic analysis does not point to a marked difference in articulatory complexity between stimulus sets. While more instances of assimilation, that is, readily co-articulated coda-onset consonants (e.g. *phone* → *ball*), appear likely for HF than LF words, this is offset by a greater proportion of word boundaries requiring major place changes in articulation in the HF set.

The stimuli were digitally recorded in a native Australian female voice, and using the ProTools LE software on a G4 Macintosh computer converted to sound files. Experimental sessions were run on an IBM compatible PC using prepared script files that were loaded into SuperLab Pro v. 2.0.4 (Cedrus Corporation, 2003). Amplification of the sound files was through an external speaker attached to the computer.

Each script file contained 64, six word trials testing four conditions (pure HF lists, pure LF lists, and half lists beginning with either HF or LF items). Half lists were constructed so that the first and second halves of the list were drawn from different stimulus sets (HHHLLL or LLLHHH). In these trials, each item of the LF and HF word sets was presented twice; once in a pure list and once in half list formats. Allocation of items to serial positions was random within the constraints of each list format, and the presentation of individual trials by condition was randomised.

Procedure

All participants were tested individually. The total time to complete the experiment was

approximately 45 minutes. Initiation of each trial occurred when the participant pressed the spacebar of the keyboard and the stimuli were presented at a rate of one word per second. After the presentation of the sixth word, a recall prompt appeared (“?????”) on the screen indicating that participants should commence the recall phase. Spoken recall was according to strict serial recall criteria, that is, (i) words were recalled in order of presentation; (ii) if a word could not be recalled the participant would indicate by saying ‘blank’; and (iii) previous items were not to be recalled after moving on to successive items in the list. Four practice trials were used to familiarise participants with the task requirements before the commencement of the experimental phase.

Results

For each participant and each condition, recall was scored and collapsed across trials to provide the number of correct items by serial position. Correct recall was scored using a strict criterion, namely responses were considered correct if a word was recalled in the position corresponding to its serial order in presentation. The mean number of correctly recalled items by serial position and condition is shown in Figure 1.

Serial recall. The criterion for significance and statistical tests were the same as for Experiment 1. The 4 x 6 (list type x serial position) repeated measures ANOVA identified that list type, $F(3,105) = 62.13, p < .001, MSE = .021, \eta_p^2 = .640$, serial position, $F(5,175) = 120.84, p < .001, MSE = .086, \eta_p^2 = .775$, and the list type by position interaction, $F(5,525) = 6.20, p < .001, MSE = .016, \eta_p^2 = .150$, were significant. Lists starting with HF items generated better performance than those beginning with LF items. In a strict sense, the serial recall curves for half lists were observed to crossover at approximately the fifth serial position, while the pure lists showed the typical difference between HF and LF items across

all serial positions. However the frequency effect for half lists at the fourth serial position was clearly influenced by the transition from HF to LF words, and vice versa.

The interaction was explored further by conducting individual analyses for each list format (pure and half lists). In the case of pure lists all results were significant, namely frequency, $F(1,35) = 123.76, p < .001, MSE = .028, \eta_p^2 = .780$, serial position, $F(5,175) = 86.08, p < .001, MSE = .056, \eta_p^2 = .711$, and the list type by position interaction, $F(5,175) = 3.56, p = .004, MSE = .010, \eta_p^2 = .092$. In this instance, the interaction was driven by a ceiling effect operating for HF items in the first serial position. A 2 x 5 ANOVA on the last five serial positions identified that while the main effects of list type, $F(1,35) = 127.08, p < .001, MSE = .028, \eta_p^2 = .784$, and serial position, $F(4,140) = 50.29, p < .001, MSE = .048, \eta_p^2 = .590$, remained significant, the interaction disappeared, $F(4, 140) = 0.53, p = .715, MSE = .012, \eta_p^2 = .015$.

The half lists also produced significant results for list type, $F(1, 35) = 29.85, p < .001, MSE = .014, \eta_p^2 = .460$, serial position, $F(5, 175) = 118.70, p < .001, MSE = .037, \eta_p^2 = .772$, and a significant interaction, $F(5, 175) = 13.33, p < .001, MSE = .017, \eta_p^2 = .276$. These results reflected the superior list-wide recall for lists beginning with HF words and the convergence of performance of half lists at approximately the fifth serial position.

A final analysis on the serial recall data distilled the nature of the frequency effect for pure and half lists by collapsing recall across the first and final three list positions. A 4 x 2 (list type x list half) ANOVA on this data revealed significant effects of list type, $F(3,105) = 62.13, p < .001, MSE = .063, \eta_p^2 = .640$, and list half $F(1,35) = 152.03, p < .001, MSE = .344, \eta_p^2 = .813$, and a significant list type by list half interaction $F(3,105) = 16.82, p < .001, MSE = .037, \eta_p^2 = .325$. The interaction was investigated by a series of Bonferroni-adjusted comparisons focused primarily on the behaviour of half lists: half lists commencing with HF items were significantly better recalled than half lists starting with LF items across early,

$t(35) = 6.87, p < .001, d = 1.125$, but not late list items, $t(35) = -0.66, p = .513, d = .052$, while recall for half lists commencing with HF items were recalled as well as pure HF lists for early serial positions, $t(35) = 0.32, p = .750, d = .061$, and half lists commencing with LF items were recalled as well as pure LF lists for the first half of the list, $t(35) = -.85, p = .399, d = .120$. Accordingly, in the current work, the recall of half lists can be characterised as mimicking recall for pure lists in the first three serial positions but producing equivalent recall in the final positions that is intermediate with respect to the envelope produced by the pure lists.

Item analysis. The responses for each trial were classified as either correct (the correct item in the correct serial position), an order error (the item recalled corresponded to a word that was presented elsewhere in the trial) or an item error (the response did not match any item presented in the trial, or was a repetition of a previously recalled list item). The proportion of errors of each classification, for each item type in each list type, collapsed across serial position and participants is given in Table 2, together with the proportion of items correct. Total order errors were conditionalized by dividing the total number of order errors by the number of items correctly recalled regardless of order. This was done to prevent the confounding of order errors with the level of item memory observed for each stimulus-condition type (Murdock, 1976; Poirier & Saint-Aubin, 1996; Saint-Aubin & Poirier, 1999). Conditionalized order errors identified that .100 of the HF items in pure lists were recalled in the wrong position, while LF items in pure lists were recalled out of position on .102 of occasions. The equivalent error values for HF and LF items in half lists respectively were .106 and .112. When submitted to a 2 x 2 (frequency x list format) repeated measures ANOVA, neither frequency, $F(1,35) = 0.27, p = .608, MSE = .002, \eta_p^2 = .008$, list type, $F(1,35) = 0.76, p = .389, MSE = .003, \eta_p^2 = .021$, nor the interaction $F(1,35) = 0.05, p = .823, MSE = .002, \eta_p^2 = .001$, were significant.

The corresponding analysis for total item errors identified that there was a significant main effect of frequency, $F(1,35) = 253.07, p < .001, MSE = .003, \eta_p^2 = .879$, no effect of list format, $F(1,35) = 0.15, p = .705, MSE = .004, \eta_p^2 = .004$, but a significant interaction, $F(1,35) = 55.29, p < .001, MSE = .003, \eta_p^2 = .612$. High frequency words produced fewer item errors than LF words, but this difference was smaller for half lists.

Discussion

This experiment was designed to contrast the recall of pure lists with half lists, and to provide a comparison with the recall of alternating lists observed by Hulme et al. (2003).² When items were presented in pure lists an HF advantage in recall across serial positions was found. Although a frequency by serial position interaction was present, this was attributable to the presence of a ceiling effect operating on the first serial position. Therefore, the results for pure lists are comparable to those found in Hulme et al. (2003, Experiment 2).

The correct recall of mixed lists across serial positions was greater for lists beginning with HF than LF words. The half lists were found to follow the envelope of the first three positions of the pure lists; the differences between half lists beginning with HF and LF words were significant across these positions. In contrast, at and beyond the change in item type in half lists, recall was intermediate with respect to the pure serial recall curves. Furthermore, in the second half of the lists, recall for half lists was not different in any serial position, although the non-significant trend identified that the curves crossed over at the fifth position and separated again at the recency position.

² To ensure that the use of the current stimuli produced comparable results with alternating lists, they have been used to replicate Hulme et al.'s (2003) Experiment 2 in our laboratory although it is not reported here.

It is possible that an attenuated frequency effect is present at the recency position in half lists, although it is not obvious from this data whether such an effect would be a product of item-to-item association or a function of the terminal position in the list. Furthermore, the lack of difference between half lists in positions four and five occurs where performance is lowest, and so the influence of a floor effect in this region of the serial recall curve is possible. To pursue these issues further, Miller and Roodenrys (in prep) tested lists where half-list-like three-item sequences were embedded into surrounds comprising items of opposite frequency (i.e. HHLLLH and LLHHHL). Importantly, sequences began at position three, away from the lowest performance levels (in this experiment recall accuracy for position three was approximately 60%). The abolition of the frequency effect in the recall of these lists coincided with the commencement of the sequence in position three and was maintained for the three-item sequence, suggesting that the lack of difference observed in Experiment 1 is not a product of a floor effect, and the emergence of any effect at recency is likely to be driven by end-of-list influences.

In line with previous research, order errors were made at the same rate for HF and LF words (Hulme et al., 2003; Poirier & Saint-Aubin, 1996; Stuart & Hulme, 2000), and more item errors occurred for LF than HF words although this difference was attenuated in half lists. This suggests that the frequency effect in serial recall arises from how composition and arrangement of items impacts memory for HF and LF items respectively, assuming that the proportion of ‘theoretical’ order errors amongst item errors for HF and LF words is constant, namely the proportion of omitted items that would otherwise be order errors, remains the same.

Experiment 1 therefore identified that the serial recall pattern produced for half lists is distinctly different to the recall of alternating lists; that is, variation in item arrangement provided evidence of a directionally sensitive component to the frequency effect. However,

this influence appears not to be effective after the change in item type within half lists (serial position four), as no difference between HF and LF items was observed for half lists in the last three serial positions.

These results are problematic for a non-directional and non-specific associative account of redintegration (Hulme et al., 2003) as according to this view the inter-item associations of all list items contribute to the level of supportive activation that assists in the reconstruction of all items. Lists with the same composition (e.g. alternating and half lists) should produce similar serial recall curves. It would appear that, for the first half of the list at least, item-to-item association as defined by presentation order is important.

While Hulme et al. (2003) argued against item-specific redintegration as the basis for the frequency effect other researchers have been more reluctant to discount it. Saint-Aubin and LeBlanc (2005) found that an LF isolate in an otherwise HF list was recalled as well as a HF item in a pure list, while a HF isolate inserted into an LF background was better recalled than a LF item in a pure list, but less well than a HF item in a pure list. The authors suggested that the benefits of item distinctiveness combined with the overall processing demands of item-specific encoding and redintegration could explain the frequency effect in mixed lists. They argued that HF items are more distinctive from other HF words than LF items from their LF neighbors in psychological space, although in mixed lists the relative distinctiveness of LF items may be enhanced if they are embedded in a HF surround. In addition, the processing demands of list items produce a list-level effect that combines with item distinctiveness to determine recall performance; LF items are more resource-demanding than HF words, so the greater the number of LF items in the list, the greater the relative cost to other list items in available resources for item-specific redintegration when compared to a pure HF list. Similarly, the greater the number of HF words in the list, the greater processing benefit to other list items when compared to a pure LF list.

However, this account also falls short in explaining the results of Experiment 1. Item distinctiveness plus processing would anticipate frequency effects in both halves of the list due to the greater distinctiveness of consecutive HF than LF items. These would be attenuated in comparison to the size of effect observed with pure lists because of the item processing costs and benefits (when compared to pure HF and pure LF lists respectively). Furthermore, as the lists are balanced in terms of item composition, neither list type (HHHLLL or LLLHHH) should offer better item processing conditions than the other and the global performance for half lists beginning with HF and LF words should be equivalent. The characteristics of recall performance in Experiment 1 are clearly at odds with these predictions.

The co-articulation hypothesis is better positioned to explain the pattern of performance observed for half lists. This position argues that recall should be sensitive to the order of items as they relate to the formulation of a speech output plan. Sequences of HF items at the start of a list should facilitate more efficient co-articulation during rehearsal, the primary means of building a single sequence of output on which to base recall. Furthermore, the absence of a frequency effect in the second half of the list might reflect the impact of the items in the first half on recall; when the first half items are HF, then benefits realised early in the speech plan might allow the later LF items to be better recalled than in pure LF lists, while lists commencing with LF items would produce speech output plans that would disadvantage the recall of HF items in the latter serial positions, relative to pure HF lists.

Experiment 2

The aim of Experiment 2 was to investigate the importance of subvocal rehearsal in the generation of the frequency effect in half lists, given that this activity is central to the co-

articulation hypothesis. Previous research using memory span tasks has shown that a frequency effect in pure lists survives conditions where subvocal rehearsal is interrupted by articulatory suppression (AS) (Gregg et al., 1989; Tehan & Humphreys, 1988). Nonetheless, Experiment 2 examined recall performance of half lists under quiet and AS conditions to extend these findings to mixed list conditions and the serial recall paradigm where the majority of participants will encounter list lengths greater than memory span. If the frequency effect is a result of mnemonic processes and separate from subvocal articulation, then the persistence of the effect under suppression conditions should be observed, in spite of an anticipated reduction in the overall level of recall. However, if the co-articulation hypothesis is correct a reduction in the frequency effect under AS should be witnessed, as the inability to engage fluently in the development of a speech output plan should limit the extent that HF items can facilitate articulatory fluency. Failure to demonstrate an attenuated effect suggests, at the very least, the frequency effect occurs at a more central level of speech planning, if it is not a product of memory processes.

For Experiment 2, a steady-state instantiation of AS was chosen; participants were asked to repeat ‘the’ throughout the presentation of list items. This contrasts with changing-state AS where a sequence of syllables with changing phonemic content, e.g. a series of letter or digit names, is used. Steady-state AS does not harm verbal recall as much as changing-state AS (Macken & Jones, 1995; Packiam Alloway, Kerr, & Langheinrich, 2010; Toppino & Pisegna, 2005), although the latter is considered to be a superior means of eliminating rehearsal. However, given the real possibility that AS could reduce overall performance levels to a point where floor effects were present, and this would make interpretation of the results difficult, the use of steady-state AS in this case was viewed as the better choice. It would still engage participants in a secondary task that would disrupt the sequencing of motor-action planning, while limiting effects of additional factors that might otherwise

reduce the size of the frequency effect and erroneously implicate co-articulation. Steady-state AS therefore provides a strong test of the co-articulation hypothesis given the stated importance of rehearsal in achieving co-articulation benefits and the reliance on articulation duration measures to support it (Woodward et al., 2008).

Method

Participants

Thirty-six undergraduate University of Wollongong students (27 females, 13 males) participated in this experiment for course credit. All participants were native Australian-English speakers and their mean age was 21.2 years ($SD = 4.3$ years).

Materials

The same stimuli that were used in Experiment 1 were employed in Experiment 2. All lists presented in this experiment conformed to half list formats. Script files for each experimental session contained two sets of 32, six word trials testing the two half list conditions (16 trials of each list type). Each item appeared in each set of trials once. Allocation of items to serial positions was random within the constraints of each list type, and the presentation of individual trials by condition in each set was randomised. One set of trials was used to test recall under quiet conditions, where participants were allowed to rehearse freely. In the AS condition, participants were instructed to repeat ‘the’ throughout the presentation of the list, at a rate of approximately 2 repetitions per second. The order of quiet and AS conditions was counterbalanced across participants.

Procedure

The procedure for administering the experiment was similar to the procedure for Experiment 1, with the exception that the practice trials included two examples of quiet and

AS trials each. The experimenter monitored the participant's adherence to AS instructions. Deviations from the required suppression rate were corrected promptly.

Results

The scoring methods used in Experiment 1 were applied to this data. The mean number of correctly recalled items by serial position and condition are shown in Figure 2. It is readily apparent from this figure that despite a noticeable reduction in recall performance in the suppression conditions, the pairs of half lists produce similar patterns across serial positions.

Serial recall. The criterion for significance and statistical tests were the same as for Experiment 1. A 2 x 2 x 6 (suppression x list type x serial position) repeated measures ANOVA identified that all three main effects were significant: suppression - $F(1,35) = 130.06, p < .001, MSE = .056, \eta_p^2 = .788$, list type - $F(1,35) = 27.49, p < .001, MSE = .020, \eta_p^2 = .440$, and serial position - $F(5,175) = 143.74, p < .001, MSE = .055, \eta_p^2 = .804$. Quiet conditions produced better recall than AS conditions, performance on the half lists replicated the pattern found in Experiment 1 where lists beginning with HF items were better recalled than lists beginning with LF items (as a consequence of a frequency effect in the first but not second half of the lists), and the serial position curve showed typical primacy and recency effects.

The two-way interaction between suppression and list type was not significant, $F(1,35) = 0.07, p = .792, MSE = .026, \eta_p^2 = .002$; the frequency effect between half lists did not vary as a function of suppression condition. However the suppression by serial position interaction was found to be significant, $F(5,175) = 6.64, p < .001, MSE = .023, \eta_p^2 = .160$, revealing that recall in the AS conditions fell away more rapidly than in the quiet conditions. Finally, the

list type by serial position interaction was also significant, $F(5,175) = 21.51, p < .001, MSE = .010, \eta_p^2 = .381$, reflecting the elimination of the effect in the later serial positions.

Importantly, the critical interaction between suppression, list type and serial position was not significant, $F(5,175) = 1.37, p = .239, MSE = .009, \eta_p^2 = .038$. Therefore, the difference between HL lists and LH lists at each serial position did not vary as a function of suppression. As in Experiment 1, HF items in the final position were, at least in descriptive terms, better recalled than LF words.

Item analysis. In line with the error analyses presented for Experiment 1, the proportions of order and item errors, together with the proportion of correct recall, were determined for each frequency by list format combination, and are given in Table 3. The rate of conditionalized order errors was .155 for the HF items in half lists recalled under quiet conditions, while it was .145 for LF items in quiet conditions. The equivalent error values for HF and LF items in half lists in AS conditions were .235 and .222, respectively. A 2 x 2 (suppression x frequency) repeated measures ANOVA identified that suppression, $F(1,35) = 49.30, p < .001, MSE = .004, \eta_p^2 = .585$, but not frequency, $F(1,35) = 1.12, p = .296, MSE = .004, \eta_p^2 = .031$, nor the interaction $F(1,35) = 0.10, p = .752, MSE = .002, \eta_p^2 = .003$, impacted the level of order errors observed. More order errors occurred under AS than quiet conditions.

A second 2 x 2 ANOVA for total item errors identified significant main effects of suppression, $F(1,35) = 155.24, p < .001, MSE = .006, \eta_p^2 = .816$, and frequency, $F(1,35) = 74.82, p < .001, MSE = .002, \eta_p^2 = .681$, but a non-significant interaction, $F(1,35) = 0.29, p = .595, MSE = .002, \eta_p^2 = .008, ns$. Greater levels of item errors were associated with LF words and recall under AS.

Discussion

Experiment 2 sought to determine whether the recall pattern for half lists under conditions where participants were free to rehearse was altered by the introduction of AS and the interruption to subvocal rehearsal. The reduction in recall under suppression conditions in this experiment is comparable to that found in previous work using the same form of suppression (e.g. D.J. Murray, 1968). Furthermore, the results of this experiment demonstrate that apart from reducing the overall level of recall and increasing the rate of change in recall in the primacy positions, AS does not change the relationship between half lists or the nature of the frequency effect across serial positions. Consequently, results are consistent with the word span results of Tehan and Humphreys (1988) who found no reduction in the frequency effect with pure lists under suppression.

Therefore, if the frequency effect can withstand the forced interruption to subvocal rehearsal that AS produces then it is likely that the effect is determined by processes prior to, or independent of, rehearsal. Accordingly, the results discount the co-articulation hypothesis as stated and suggest that subvocal rehearsal is not critical to the formation of the frequency effect.

The results of item analyses are consistent with those found in Experiment 1 and other work (e.g. Hulme et al., 2003). That is, the frequency effect is a consequence of differing levels of memory for HF and LF items and not a result of variation in order errors.

General Discussion

The results from these experiments clearly show that in their current form neither of the candidate frameworks can explain the key observations of this research; a changing frequency effect with serial position in half lists, the elimination of the effect in the second half of half lists, and the inability of AS to eliminate or reduce the effect. The co-articulation

hypothesis places the locus of the frequency effect with speech motor planning, but highlights the rehearsal of list items as a key factor in the formation of a plan to direct output as a single utterance (Woodward et al., 2008). According to this view, language skills and speech habits reflecting linguistic familiarity moderate the capacity of the output planning mechanism to recreate the sequence of presented items. Familiarity of the transitions between items, argued to be correlated with word frequency, determines the co-articulatory fluency of the output sequence. Therefore, by this account, differences in co-articulatory fluency are responsible for the frequency effect in serial recall performance.

The co-articulation hypothesis (Woodward et al., 2008) is drawn from the perceptual-gestural account of STM (Jones et al., 2006; Hughes et al., 2009; A. Murray & Jones, 2002). This position argues that serial recall performance does not reflect the use of specialized memory stores and associated operations but rather the exploitation of perceptual and articulatory properties in a sequence of unrelated items that assist in maintaining item order. Under these conditions, stimuli form a perceptual stream which feeds directly into output planning, and makes use of speech motor processes to produce a single motor plan of articulatory gestures for output. That is, the perceptual process feeds directly into the output planning process and no reference to non-articulatory (i.e. phonological) knowledge is made. Rehearsal represents the action of the output plan, not the refreshing of short-term representations contained in a mnemonic store.

This approach is capable of explaining the results of Experiment 1. As described more fully above, word frequency affects familiarity with the transitions at word boundaries, and therefore fluency in speech output planning. The abolition of the frequency effect in the second half may be explained by the additional assumption of a limited resource and flow on effects from the first half of the list. The results of Experiment 2 are more troublesome for the co-articulation hypothesis. The frequency effect in half lists, and in particular how it

manifests across serial positions, was not altered by conditions that required participants to engage in AS. In this experiment, as indicated by the overall effect of suppression on recall, rehearsal for target material was restricted, if not disabled. Despite this, no reduction in the frequency effect was observed. Therefore, in order for a perceptual-gestural account of STM to accommodate the results of Experiment 2, the locus of the frequency effect must necessarily occur within an articulatory process that operates prior to, and is independent of, subvocal rehearsal. Furthermore, these co-articulatory factors must not require repeated activation during pre-implementation in order to manifest at output.

The HF and LF stimuli used in the current experiments were found to be minimally different with respect to measures likely to influence co-articulation and were monosyllabic. Therefore it is possible that these experiments did not provide an opportunity for co-articulatory effects, in addition to effects with other loci, to arise. If so, it is not surprising that the frequency effect did not diminish when participants were prevented from rehearsing target material freely. However, it is equally the case that a strong interpretation of the co-articulation hypothesis, namely that the frequency effect is solely the product of co-articulation differences, is clearly incompatible with these results.

Redintegration, as defined by Hulme et al. (2003; Stuart & Hulme, 2000, 2009), also cannot explain the observed sensitivity of the frequency effect to item arrangement. Yet the results of Experiment 2 suggest the effect is likely to be mnemonic in nature. The presence of an item-to-item influence, even for early serial positions, implies some seriation in the development in a list-based network of activation that impacts recall. However, how such an effect can arise in a model where LTM representations are activated during presentation, presumably forming a mature activation network prior to the retrieval phase, poses a challenge requiring further work.

Allen and Hulme (2006) examined the frequency effect in serial recall (in pure lists) as well as a series of speech perception tasks (auditory lexical decision speed, word identification in noise), and speech production tasks (definition naming, delayed repetition, and maximal articulation rate). The results identified the speech processing measures of definition naming accuracy and word identification in noise as predictors of serial recall performance, although the latter was a weak predictor. A positive association was found between serial recall and definition naming accuracy (the generation of an item on the basis of its semantic description), as HF words were more likely than LF words to be recalled in a serial recall task and retrieved from their semantic representation.

Allen and Hulme (2006) considered the psycholinguistic model of Martin and Saffran (1997) to be compatible with the findings presented in their study. This model assumes that auditory-verbal STM is a capacity derived from acquired language knowledge. Consequently the short-term maintenance and retrieval of items is dependent on the temporary activation of LTM representations of items at multiple levels in the language processor. The processing of auditory input involves activation of the phonological representations and the spread of activation then engages the lexical and semantic representations of items. In turn, this activation provokes feedback activation to each of the preceding levels in the system. Perpetuation of this activity maintains item information until retrieval. The output pathway commences with activation of the semantic features of a word that feed forward to the lexical and phonological representations. Feedback from these layers produces a cyclic pattern that, when stabilised, will select a candidate for phonological encoding.

Serial order is maintained in this model by connections between representations in each of these layers and a sequence placeholder, that are established as a string of words is processed (Martin & Saffran, 1997). Consequently, the influence of lexico-semantic representations is concentrated on item memory. This is consistent with the small to non-

existent effects on order produced by lexico-semantic variables generally (Majerus, 2009), and is supported by the absence of frequency-based order effects in the present work.

The findings from the current experiments suggest that in the serial recall task, the act of producing a sequence of items at output might serve to create a pattern of activation within the speech motor processes that reflects the semantic association between adjacent items. This would underpin the development of an item-to-item associative influence that occurs unambiguously for early list positions in half lists. The abolition of an item-to-item effect for the second half of lists might reflect the maturation of such an activation network that occurs as the final list items are outputted. Alternatively, relative efficiencies in the output process that are determined by the attributes of early list items might be responsible; half lists that commence with a series of HF items might leave more resources available for the recall of the following LF items, and produce performance that is superior to that observed in pure LF lists. Conversely, the poorer recall of HF items in the second half of half lists, when compared to recall in pure HF lists, might point towards a reduction in the resources available due to the demands involved with recalling a series of LF items earlier in the list.

These results render the problem of developing a computational model capable of replicating the frequency effect in serial recall challenging. The effect has not been central to the design of computational models of STM (e.g. Brown, Neath, & Chater, 2007; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Farrell & Lewandowsky, 2002; Henson, 1998; Lewandowsky & Murdock, 1989; Neath, 2000; Page & Norris, 1998) because frequency does not appear to influence order memory and this is the primary problem that they aim to resolve. Although some models, utilising a variety of mechanisms, can accommodate the effect found with pure lists (e.g. Burgess & Hitch, 1999; Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, 2000, 2008; Page & Norris, 1998), none has been developed to account for the abolition of the effect in alternating lists. Moreover, the

pattern of recall found with half-lists underscores the possibility that the effect is a composite of a number of factors that contribute to varying degrees at different points in the recall of a sequence of items. In future, greater attention to modelling the roles of lexico-semantic variables will be required if computational models are to extend their explanatory capacity and encompass these effects.

In summary, the empirical data presented here echo the sentiment expressed by Hulme et al. (1997) – that the frequency effect in serial recall is complex and the identification of mechanisms responsible for its expression is unlikely to be straightforward. The results of the current experiments suggest that it is predominantly a late-stage effect that is shaped by the processes associated with output. The effect would appear to be consistent with a psycholinguistic explanation of STM, however, it remains unclear which specific point(s) along the speech production pathway are responsible for the formation of the frequency effect across different serial positions. What is clear is that the effect cannot be either a result of non-directional and non-specific activation supporting recall as the redintegration hypothesis suggests, or due to articulatory differences at word boundaries modifying the impact of rehearsal as the co-articulation hypothesis maintains.

References

- Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, 55, 64-88.
- Baayen, R.H., Piepenbrock, R., & Van Rijan, H. (1993). *The Celex Lexical Database*. Philadelphia: Linguistic data consortium, University of Pennsylvania.
- Baddeley, A.D. (1986). *Working memory*. New York: Oxford University Press.
- Brown, G.D.A., Neath, I., & Chater, N. (2007). A Temporal Ratio Model of Memory. *Psychological Review*, 114, 539-576.
- Brown, G.D.A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, 107, 127-181.
- Burgess, N., & Hitch, G.J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, 106, 551-581.
- Cedrus Corporation. (2003). *SuperLab Pro: Experimental Laboratory Software*. San Pedro, USA: Cedrus Corporation.
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505. Retrieved from http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm
- Deese, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports*, 5, 305-312.
- Deese, J. (1960). Frequency of usage and number of words in free recall: The role of association, *Psychological Reports*, 7, 337-344.
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9, 59-79.

- Farrell, S., & Lewandowsky, S. (2008). Empirical and theoretical limits on lag recency in free recall. *Psychonomic Bulletin & Review*, 15, 1236-1250.
- Gregg, V., Freedman, C.M., & Smith, D.K. (1989). Word frequency, articulatory suppression and memory span. *British Journal of Psychology*, 80, 363-374.
- Henson, R.N.A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, 36, 73-137.
- Howard, M.W., & Kahana, M.J. (2002a). A distributed representation of temporal context. *Journal of Mathematical Psychology*, 46, 269-299.
- Howard, M.W., & Kahana, M.J. (2002b). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, 46, 85-98.
- Hughes, R.W., Marsh, J.E., & Jones, D.M. (2009). Perceptual-gestural (mis)mapping in serial short-term memory: The impact of talker variability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1411-1425.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G.D.A, Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1217-1232.
- Hulme, C., Stuart, G., Brown, G.D.A., & Morin, C. (2003). High- and low-frequency words are recalled equally well in alternating lists: Evidence for associative effects in serial recall. *Journal of Memory and Language*, 49, 500-518.
- Jones, D.M., Hughes, R.W., & Macken, W.J. (2006). Perceptual organization masquerading as phonological storage. *Journal of Memory and Language*, 54, 265-281.
- Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. In B.H. Ross (Ed). *The psychology of learning and motivation: Advances in research and theory* (Vol 49). (pp. 1-48). San Diego, CA, US: Elsevier Academic Press.

- Lewandowsky, S., & Murdock, B.B. Jr. (1989). Memory for serial order. *Psychological Review*, 96, 25-57.
- Macken, W.J., & Jones, D.M. (1995). Functional characteristics of the inner voice and the inner ear: Single or double agency? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 436-448.
- Majerus, S. (2009). Verbal short-term memory and temporary activation of language representations: The importance of distinguishing item and order information. In A. Thorn and M. Page (Eds), *Interactions between short-term and long-term memory in the verbal domain* (pp. 244-276). Hove, UK: Psychology Press.
- Martin, N. & Saffran, E. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, 14, 641-682.
- Morin, C., Poirier, M., Fortin, C., & Hulme, C. (2006). Word frequency and the mixed-list paradox in immediate and delayed serial recall. *Psychonomic Bulletin & Review*, 13, 724-729.
- Mueller, S.T., Seymour, T.L., Kieras, D.E., & Meyer, D.E. (2003). Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1353-1380.
- Murdock, B.B. Jr. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General*, 105, 191-216.
- Murray, A., & Jones, D.M. (2002). Articulatory complexity at item boundaries in serial recall: The case of Welsh and English digit span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 594-598.

- Murray, D.J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, 78, 679-684.
- Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review*, 7, 403-423.
- Packiam Alloway, T., Kerr, I., Langheinrich, T. (2010). The effect of articulatory suppression and manual tapping on serial recall. *Journal of Cognitive Psychology*, 22, 297-305.
- Page, M.P.A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, 105, 761-781.
- Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology*, 50, 408-412.
- Roodenrys, S., Hulme, C., Alban, J., Ellis, A.E., & Brown, G.D.A. (1994). Effects of word frequency and age of acquisition on short-term memory span. *Memory & Cognition*, 22, 695-701.
- Saint-Aubin, J., & LeBlanc, J. (2005). Word frequency effects in immediate serial recall of pure and mixed lists: Tests of the associative link hypothesis. *Canadian Journal of Experimental Psychology*, 59, 219-227.
- Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory, and Cognition*, 21, 168-175.
- Stuart, G., & Hulme, C. (2000). The effects of word co-occurrence on short-term memory: Associative links in long-term memory affect short-term memory performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 796-802.
- Stuart, G.P., & Hulme, C. (2009). Lexical and semantic influences on immediate serial recall: A role for redintegration. In A. Thorn and M. Page (Eds), *Interactions between Short-term and Long-term Memory in the Verbal Domain* (pp. 157-176). Hove, UK: Psychology Press.

- Tehan, G., & Humphreys, M.S. (1988). Articulatory loop explanations of memory span and pronunciation rate correspondences: A cautionary note. *Bulletin of the Psychonomic Society*, 26, 293-296.
- Toppino, T.C., & Pisegna, A. (2005). Articulatory suppression and the irrelevant speech effect in short-term memory: Does the locus of suppression matter? *Psychonomic Bulletin & Review*, 12, 374-379.
- Watkins, O.C., & Watkins, M.J. (1977). Serial recall and the modality effect: Effects of word frequency. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 712-718.
- Woodward, A.J., Macken, W.J., & Jones, D.M. (2008). Linguistic familiarity in short-term memory: A role for (co-)articulatory fluency?, *Journal of Memory and Language*, 58, 48-65.
- Wright, C.E. (1979). Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory & Cognition*, 7, 411-419.

Table 1

Phonetic characteristics of word boundaries for the HF and LF word sets used in Experiments 1-2.

Phonetic characteristic	HF	LF
Manner of articulation		
Combinations with the same manner of articulation	2411 (.264)	2541 (.279)
CC combinations requiring a change in manner of articulation	6709 (.736)	6579 (.721)
Place of articulation		
CC combinations with the same place of articulation	2965 (.325)	2906 (.319)
CC combinations with a different place but assimilation likely	1198 (.131)	815 (.089)
CC combinations requiring a change of place of articulation		
Minor place change (e.g. between 2 anterior Cs)	2134 (.234)	2822 (.309)
Major place change (e.g. between anterior/posterior Cs)	2823 (.310)	2603 (.285)
Voicing		
CC combinations involving no change in voicing	4641 (.509)	4703 (.516)
CC combinations involving a change in voicing	4479 (.491)	4417 (.484)

Note. C = consonant. Proportions are in brackets.

Table 2

Mean proportion of items correctly recalled and mean proportions of different error categories as a function of item type and list type in Experiment 1

	Correct	Errors	
		Order	Item
HF in pure lists	.684 (.143)	.071 (.046)	.244 (.117)
LF in pure lists	.504 (.134)	.052 (.033)	.444 (.119)
HF in half lists	.623 (.126)	.071 (.043)	.303 (.110)
LF in half lists	.560 (.135)	.065 (.045)	.377 (.116)

Note. Standard deviations are given in brackets.

Table 3

Mean proportion of items correctly recalled and mean proportions of different error categories as a function of item type and list type in Experiment 2

	Correct	Errors	
		Order	Item
HF quiet	.608 (.170)	.101 (.068)	.300 (.143)
LF quiet	.546 (.186)	.080 (.047)	.365 (.163)
HF AS	.427 (.134)	.124 (.061)	.461 (.121)
LF AS	.359 (.126)	.094 (.034)	.534 (.121)

Note. Standard deviations are given in brackets. AS = articulatory suppression.

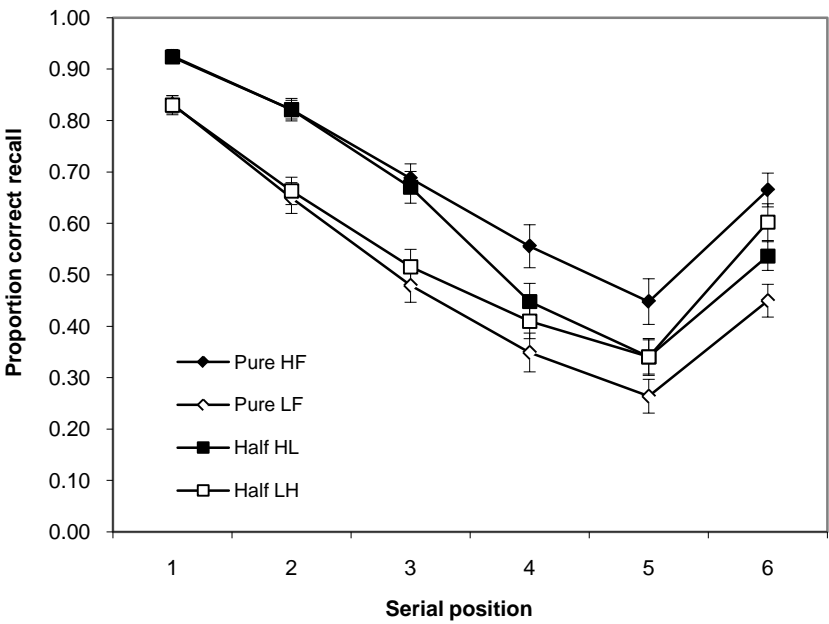


Figure 1. Serial recall of words as a function of frequency and list composition in Experiment

1. Error bars represent standard errors.

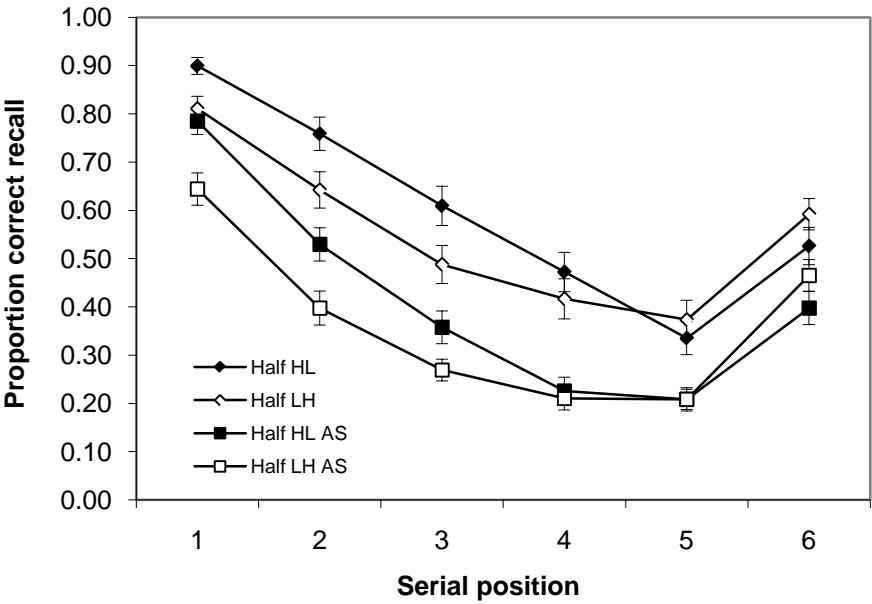


Figure 2. Serial recall of words as a function of frequency and list composition in Experiment 2. Error bars represent standard errors. AS = articulatory suppression.